Overcoming health disparities in precision medicine

Kathleen C. Barnes and Francisco M. De La Vega
Tempus Labs, Inc.
Chicago, IL 60654, USA
Email: kathleen.barnes@tempus.com; francisco.delavega@tempus.com

Carlos D. Bustamante
Galatea Bio, Inc.
Miami, FL, USA
Email: carlos.bustamante@galatea.bio

Chris R. Gignoux
University of Colorado Anschutz Medical Campus
Boulder, CO, USA
Email: hris.gignoux@cuanschutz.edu

Eimear Kenny
Ichan School of Medicine at Mount Sinai
New York, NY, USA
Email: eimear.kenny@mssm.edu

Rasika A. Mathias
Johns Hopkins School of Medicine
Baltimore, MD, USA
Email: rmathias@jhmi.edu

Bogdan Pasaniuc
University of California Los Angeles.
Los Angeles, CA, USA
Email: niuc@ucla.edu

1. Overview

Precision medicine and precision public health rely on the premise that determinants of disease incidence and differences in response to interventions can be identified and their biology understood well enough that applications to reduce risk of disease and improve treatment can be developed. However, there are well-documented racial and ethnic disparities throughout health care at the patient, provider, and healthcare system levels. These disparities are driven by a complex

© 2022 The Authors. Open Access chapter published by World Scientific Publishing Company and distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License.

interplay among social, psychosocial, lifestyle, environmental, health system, and biological determinants of health (Freedman, et al. 2021).

Inequities in genome-informed precision medicine are driven by a Eurocentric bias in genetic studies: the vast majority (86%) of genomics studies have been conducted in individuals of European descent. Eurocentric biases in genetics studies are not only inequitable, but also result in major missed scientific opportunities (Fatumo et al. 2022). As underrepresented minority populations within the United States grow to record numbers, and precision medicine is beginning to be deployed worldwide, it is increasingly important to invest in efforts to characterize, understand, and end racial and ethnic disparities in healthcare.

2. Equitable risk prediction

Despite the significant advances in disease risk prediction derived from the analysis of the large-scale data available in the UK Biobank, the underrepresentation of participants from minority and disadvantaged groups has limited the use of this data in the development of prediction models that can be generalized to diverse populations. The paper of Gu et al. (2023) proposes a transfer learning framework based on random forest models (TransRF) that can incorporate risk prediction models trained in a source population to improve the prediction performance in a target underrepresented population with limited sample size.

Polygenic risk scores (PRS) are numerical indicators of risk based on multiple genetic markers associated with a disease or trait and are derived from data from genome-wide association studies (GWAS). Research in this field has recently accelerated, and scores are available for a wide array of traits and conditions, including for conditions such as coronary artery disease, type 2 diabetes, and common cancers. However, research has shown that their performance is lower and somewhat unpredictable in non-European populations. In this volume, Machado Reyes et al. (2023) present a method called FairPRS, which is based on domain-adaptation problems in machine learning such as Invariant Risk Minimization (IRM) to obtain an ancestry-invariant PRS estimates from pre-computed PRS or GWAS summary statistics. FairPRS provides risk estimates with negligible effect of ancestral groups of the subjects, while increasing phenotype prediction accuracy, in both simulated and real data sets and showcases how machine learning methods can be applied to improve the portability of PRS.

Regarding disparities in outcome prediction, Chu et al. (2023) employ association rule mining, a technique that infers probabilistic implications from data in transactional databases, to identify the most significant risk categories for adverse pregnancy outcomes (APOs) in a dataset of over 10,000 nulliparous women, that is representative of the US population. Using this method, they find that the effects of age and body mass index have major yet differential effects on the risk of APOs and the observed racial/ethnic disparities. This work shows that association rule mining could be a powerful method to explore inequities in clinical datasets.

3. Pharmacoequity

While the growing body of pharmacogenomics research has significant potential for guiding treatment decisions, the persistent heterogeneity of observed treatment responses in many clinical situations suggests that additional genetic and other biologic factors may contribute to the success or failure of a given treatment approach in individuals of different racial and ethnic backgrounds. Pharmacogenomic studies have long neglected to collect data from African Americans, Hispanics/Latinos and other ethnicities, preventing an understanding of the role of ancestry in pharmacoequity. Yang et al. (2023) make some progress in this subject by analyzing the role of both global and local ancestry on measures of response to clopidogrel therapy in a cohort of 167 African American patients. They find that local ancestry at the transcription start site of three relevant genes as well as ancestry-adjusted association with variants in another gene help to explain the variability in drug response seen in African Americans.

The widespread availability of antiretroviral therapies (ART) for HIV-1 have generated considerable interest in understanding the pharmacogenomics of ART. In some individuals, ART has been associated with excessive weight gain, which disproportionately affects women of African ancestry. The paper of Keat et al. (2023) explored whether a multi-ancestry PRS for body mass index (BMI) can achieve high cross-ancestry performance for predicting baseline BMI in diverse, prospective ART clinical trials. They show that the BMI PRS explained ~5%-7% of variability in baseline BMI, with high performance in both European and African genetic ancestry groups, but that this score was not associated with the change in BMI on ART. This study thus argues against a shared genetic predisposition for baseline BMI and ART-associated weight gain.

4. Race, genetic ancestry, and population structure

A challenge in precision medicine is the continued use of “race”— a categorization based on common physical characteristics — and “ethnicity” — a categorization based on shared cultural traits — in medicine, which has become a matter of intense debate. A key element of genome-informed precision medicine is the accurate assessment and utilization of ancestry to understand its impact on disease susceptibility and the outcomes of therapies. Genomics can capture ancestry in a more precise way, allowing genetic influences to be teased apart from the impact of social and environmental factors. Understanding shared genetic ancestry and defining genetically related subpopulations can help us better understand disease susceptibilities and health disparities. Along this topic, the work of Chaichoompu et al. (2023) presents improvements in an unsupervised method, IPCAPS, to identify population substructure guided by genetic similarity. This method could be particularly useful for populations in geographically confined regions, where IPCAPS was shown to detect meaningful subgroups, which are otherwise hard to detect with classic methods such as PCA or ADMIXTURE. These subgroups can be carried downstream in population or disease association analysis instead of race/ethnicity and could prove useful in precision medicine.

5. Conclusion

The heightened impact of COVID-19 on medically underserved populations and enhanced focus on social justice issues has highlighted the need to better address health disparities in a meaningful way. New computational and statistical methods are needed to assess, counteract, and overcome health disparities in healthcare. While there is much more work to be done, we believe the work presented in this session showcases advances that will be helpful to the goal of overcoming health disparities in precision medicine.

Acknowledgments

We thank the anonymous reviewers that helped in the peer review process of the submissions to this session.

References

Chaichoompu K., Wilantho A., Wangkumhang P., TongsimaS., Cavadas B., Pereira L., and Kristel Van Steen (2023) Fine-scale subpopulation detection via SNP-based unsupervised method: A case study on the 1000 Genomes Project Resources. In. Pacific Symposium on Biocomputing 2023.

Chu H., Ramola R., Jain S., Haas D.M., Natarajan S., and Radivojac P. (2023) Using Association Rules to Understand the Risk of Adverse Pregnancy Outcomes in a Diverse Population. In. Pacific Symposium on Biocomputing 2023.

Fatumo S., Chikowore T., Choudhury A., Ayub M., Martin A.R., Kuchenbaecker K. (2022) A roadmap to increase diversity in genomic studies. Nat Med. 2022 Feb;28(2):243-250.

Freedman J.A., Abo M.A., Allen T.A., Piwarski S.A., Wegermann K., and Patierno S.R. (2021) Biological Aspects of Cancer Health Disparities. Ann. Rev. Med. 72:229-241

Gu T., Han Y., and Duan R. (2023) A transfer learning approach based on random forest with application to breast cancer prediction in underrepresented populations. In. Pacific Symposium on Biocomputing 2023.

Keat K., Hui D., Xiao B., Bradford Y, Cindi Z., Daar E.S., Gulick R, Riddler S.A., Sinxadi P., Haas D.W., and Ritchie M.D. (2023) Leveraging Multi-Ancestry Polygenic Risk Scores for Body Mass Index to Predict Antiretroviral Therapy-Induced Weight Gain. In. Pacific Symposium on Biocomputing 2023.

Machado Reyes D., Bose A., Karavani E., and Parida L. (2023) FairPRS: a fairness framework for Polygenic Risk Scores. In. Pacific Symposium on Biocomputing 2023.

Yang G., Alarcon C, Friedman P., Gong L., Klein T., O'Brien T., Nutescu E.A., Tuck M., Meltzer D., Perera M.A. (2023) The Role of Global and Local Ancestry on Clopidogrel Response in African Americans. In. Pacific Symposium on Biocomputing 2023.